PRML Assignment - 4
Submission Deadline: April 4, 2025

Instructions:
- Use Python and libraries like NumPy, Pandas, Scikit-learn, and Matplotlib/Seaborn/Plotly for implementation and visualization.

Submit:
- A Jupyter Notebook or Python script with well-commented code.
- A PDF report (1–2 pages) summarizing:
- Steps followed
- Visualization plots
- Comparison between PCA and LDA
- Ensure the code is clean, reproducible, and includes visualizations.

Question 1: Use the Iris Dataset or any high-dimensional dataset (e.g., Wine, Breast Cancer from sklearn.datasets).
1) Normalize the dataset and compute the covariance matrix.
2) Calculate eigenvalues and eigenvectors and explain the proportion of variance captured by the top 2 principal components.
3) Project the original data onto the top 2 principal components.
4) Visualize the projected data in 2D and interpret the separability of classes after PCA.

Question 2: Use the same dataset as in Question 1.
1. Implement LDA to reduce dimensions to 1D or 2D (depending on number of classes).
2. Compare and contrast the assumptions and goals of PCA and LDA in your report.
3. Project the original data onto the LDA components.
4. Visualize the LDA-projected data and discuss how class separability improves over PCA.